This chapter details the method of TEC measurement using GPS systems, which can be categorized as a typical multi-carrier multi-coded spread spectrum system. The major advantage and disadvantage of such a system is also brought out.

6.1 Introduction

A satellite navigation or SATNAV system is a system of satellites that provide autonomous geo-spatial positioning with global coverage. It allows small electronic receivers to determine their location (longitude, latitude, and altitude) to within a few metres using time signals transmitted from satellites along line-of-sight. Receivers calculate the precise time as well as position, which can be used as a reference for scientific experiments. The first satellite navigation system was Transit, a system deployed by the US military in the 1960s, the details of which have been addressed in Chapter 4.

A satellite navigation system with global coverage is termed a global navigation satellite system or GNSS. Global coverage for such a system is generally achieved by a satellite constellation of 20–30 Medium Earth Orbit (MEO) satellites spread between several orbital planes. The actual systems vary, but use orbit inclinations of >50° and orbital periods of roughly twelve hours, orbiting at an altitude of about 20,000 kilometres. The first system in this category is the NAVSTAR (Navigation System with Timing And Ranging), popularly known as the Global Positioning System (GPS). It was developed by the U.S. Department of Defence (DoD) to provide precise estimates of position, velocity, and time to users worldwide. The DoD approved the basic architecture of the system in 1973, the first satellite was launched in 1978, and the system was declared operational to public in 1996.
Today, GPS serves over twenty million users with a breath-taking variety of applications built on these basic capabilities, far out-reaching the original expectations of the system designers in the 1970s. The utilization of GPS within modern infrastructure has been so wide-ranging and prolific that it has also revolutionized defense force operations. An extensive literature on GPS is presently available. The principles of GPS are described in detail by Wells et al [1987], Kleusberg and Teunissen [1996], Parkinson et al [1996], Leick [2004], and Hoffmann-Wellenhoff et al [1994]. Theoretical and everyday practical information about GPS can be accessed via Langley [1996]. Specific GPS-related scientific and technical issues are covered in the Innovation column of the monthly GPS World magazine.

A GPS user can typically estimate location with an accuracy of better than 10 meters, and time to better than 100 nanoseconds. This is the stand-alone capability of GPS, because it does not require the user to install any local infrastructure. However, the accuracy of the location or navigation measurement is dependent on an accurate estimate of the ionospheric delay, a function of the TEC. The data received from GPS receivers are also used to study TEC. Literature survey suggests that a good amount of work has been undertaken to address in this regard, though most of the research have been for mid-latitude regions. For the Indian low-latitude region, limited work only has been reported so far, which underlines the importance of the present study.

6.2 An overview of the different GNSS constellations

As of October 2011, only the United States NAVSTAR Global Positioning System (GPS) and the Russian GLONASS are the fully globally operational GNSSs. China is in the process of expanding its regional Beidou navigation system into the global COMPASS navigation system by 2020. The European Union's Galileo positioning system is a GNSS in initial deployment phase, scheduled to be fully operational by 2020 at the earliest. Several countries including France, Japan and India are in the process of developing regional navigation systems. A brief description of some of these systems for which details are presently available is given in Table 6.1 below.
Table 6.1 Comparison of global navigation systems (upto 2012)

<table>
<thead>
<tr>
<th>System</th>
<th>GPS</th>
<th>GLONASS</th>
<th>COMPASS</th>
<th>Galileo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Political entity</td>
<td>United States</td>
<td>Russia</td>
<td>China</td>
<td>European Union</td>
</tr>
<tr>
<td>Coding</td>
<td>CDMA</td>
<td>FDMA/CDMA</td>
<td>CDMA</td>
<td>CDMA</td>
</tr>
<tr>
<td>Encoding</td>
<td>BPSK</td>
<td>BPSK</td>
<td>BOC</td>
<td>BOC</td>
</tr>
<tr>
<td>Orbital height</td>
<td>20,180 km</td>
<td>19,130 km</td>
<td>21,150 km</td>
<td>23,220 km</td>
</tr>
<tr>
<td>Inclination</td>
<td>55</td>
<td>64.8</td>
<td>56.5</td>
<td>56</td>
</tr>
<tr>
<td>No. of orbital planes</td>
<td>6</td>
<td>3</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Period</td>
<td>11.97 hours (11h 58m)</td>
<td>11.26 hours (11h 16m)</td>
<td>12.63 hours (12h 38m)</td>
<td>14.08 hours (14h 5m)</td>
</tr>
<tr>
<td>Evolution per sidereal day</td>
<td>2</td>
<td>17/8</td>
<td>36/19</td>
<td>17/10</td>
</tr>
<tr>
<td>Number of satellites</td>
<td>At least 24</td>
<td>31. Includes 24 operational, 1 in preparation, 2 on maintenance, 3 reserve and 1 on tests</td>
<td>5 geostationary orbit (GEO) satellites, 30 medium Earth orbit (MEO) satellites</td>
<td>2 test bed satellites in orbit, 22 operational satellites budgeted</td>
</tr>
<tr>
<td>Frequency</td>
<td>1.57542 GHz (L1 signal), 1.2276 GHz (L2 signal)</td>
<td>Around 1.602 GHz (SP), Around 1.246 GHz (SP)</td>
<td>1.561098 GHz (B1), 1.589742 GHz (B1-2), 1.20714 GHz (B2), 1.26852 GHz (B3)</td>
<td>1.164–1.215 GHz (E5a and E5b), 1.260–1.300 GHz (E6), 1.559–1.592 GHz (E2-L1-E11)</td>
</tr>
<tr>
<td>Status</td>
<td>Operational, CDMA in preparation</td>
<td>Operational</td>
<td>10 satellites operational, 25 additional satellites planned</td>
<td>In preparation</td>
</tr>
</tbody>
</table>

All the four systems, GPS, GLONASS, COMPASS and Galileo, mentioned above are based on direct sequence spread spectrum (DSSS) techniques. The GPS is a constellation of satellites in which some are always above the horizon (5-8 satellites always available) at any location on Earth so they are always visible. As indicated in the Table above, the GPS constellation employs 24 spacecraft in 20,200 km circular
orbits inclined at 55 degrees. These spacecrafts are placed in 6 orbital planes with four operational satellites in each plane. The GPS satellite generations are divided into four groups: I, II and IIA, IIR and IIF. The Block IIA is termed as the second generation and IIR the third generation. Block IIF is planned for replenishment with more satellite life-time and is expected to be launched in the near future. The primary difference in the various satellite generations are accuracy and maximum number of days without contact from monitoring and control stations i.e., without ground corrections and accuracy degradation.

The GLONASS constellation also has 24 satellites in circular orbits inclined at 64.8 degrees. GLONASS’ orbit makes it especially suited for usage in high latitudes (north or south), where getting a GPS signal can be problematic. The spacecrafts are placed in 3 orbits, with each orbital plane containing eight equally spaced satellites, so that one of the satellites will be at the same spot in the sky at the same sidereal time each day. The COMPASS & Galileo constellations have almost the same inclination as GPS. COMPASS has 5 satellites in geostationary orbit apart from 30 MEO satellites in six orbital planes while Galileo envisages 24 satellites in 3 orbital planes.

A comparison of the signals and their bands for L1 and L2 for the various GNSS configurations (both existing and new) is shown in figures 6.1 (a) and 6.1 (b), as given by Sergey Revnivykh in his presentation to International GNSS Committee, 2012. It is seen that the signals for L1 frequency are centered on approximately the same band for GPS, Galileo and Compass systems (1560-1590 MHz), while GLONASS signals are in 1590-1610 MHz band. It can also be seen that the L1 signals are almost identical for Galileo and Compass systems and the multiple access technique is changed from FDMA to CDMA for the newer GLONASS systems. Unlike the L1 band, the L2 signal bands are different for all the satellite configurations as generally this frequency is modulated by the precision code.
Every GPS satellite transmits the L1 frequency centered at 1576.42 MHz and L2 frequency centered at 1227.6 MHz. L1 is modulated first with C/A (Coarse/Acquisition) code, having a chipping rate of 1.023 MHz and code length of 1023 bits, giving it a refresh rate of ~1 msec. L1 is also modulated with a 50 Hz navigation message, which provides details on GPS satellite orbits and clock
corrections. The Precise (P) code modulates both the L1 and L2 carriers and has a chipping rate of 10.23 MHz.

Each GLONASS satellite transmits on slightly different L1 and L2 frequencies, with P-code on both L1 and L2, and with C/A code, at present, only on L1. GLONASS-M satellites reportedly transmit the C/A code on L2 as detailed in the *Interface Control Document of GLONASS [1998]*. Unlike GPS, all GLONASS satellites transmit the same code at different frequencies. They derive signal timing and frequencies from one of three on-board Cesium atomic clocks operating at 5 MHz.

More detailed descriptions of the GLONASS system can be found in *Langley [1997d], Parkinson et al [1996], Kleusberg and Teunissen [1996], Kaplan [1996], Hoffmann-Wellenhoff et al [1994]* among various others also.

### 6.3 Signal structure of popular GNSS systems

In order to understand the details of the transmitted signals, the signal structure of the popular GPS and GLONASS systems, studied as part of this work, is detailed below.

#### 6.3.1 GLONASS signal structure

GLONASS satellites transmit two types of signals: a standard precision (SP) signal and an obfuscated high precision (HP) signal. The signals use DSSS encoding and binary phase-shift keying (BPSK) modulation. All GLONASS satellites transmit the same code as their SP signal; however each transmits on a different frequency using a 15-channel frequency division multiple access (FDMA) technique spanning either side from 1602.0 MHz, known as the L1 band. The centre frequency of each satellite is $1602 \text{ MHz} + n \times 0.5625 \text{ MHz}$, where $n$ is a satellite's frequency channel number ($n=−7,−6,−5,...,0,...,6$, previously $n=0,...,13$). It means that satellites transmit signals on their own frequency, separated by multiples of 0.5625 MHz or 562.5 kHz, from the frequency of other satellites. Signals are transmitted in a 38° cone, using right-hand circular polarization, at an EIRP between 25 to 27 dBW (316 to 500 watts). Thus GLONASS uses one type of multi-carrier modulation. Also, it is observed that the 24-satellite constellation is accommodated with only 15 channels by using
identical frequency channels to support antipodal (opposite side of planet in orbit) satellite pairs, as these satellites will never be in view of an earth-based user at the same time, as shown in figure 6.2.

The HP signal (L2) is broadcast in phase quadrature with the SP signal, effectively sharing the same carrier wave as the SP signal, but with a ten-times-higher bandwidth than the SP signal. The L2 signals use the same FDMA as the L1 band signals, but transmit straddling 1246 MHz with the centre frequency determined by the equation $1246 \text{ MHz} + n \times 0.4375 \text{ MHz}$, where $n$ spans the same range as for L1.

### 6.3.2 GPS signal structure

GPS satellites transmit two coherent frequencies in the L-band at $f_1 = 1576.4 \text{ MHz}$ (L1) and at $f_2 = 1227.6 \text{ MHz}$ (L2). The L1 frequency is modulated by a public Coarse/Acquisition code (C/A) with an effective wavelength of 300 m. Both carrier frequencies are modulated by a precise code (P or Y) with an effective wavelength of approximately 30 m. The basic frequency structures of C/A code and P-code are shown separately as figures 6.3 (a) and 6.3 (b) and combined structure is shown in figure 6.3 (c).
In the combined signal, P-code is written 90 degrees out of phase with the C/A code (quadrature). Although all satellites transmit at the same frequency the code
differences allow them to be separated. It also means that we can track the satellites knowing the C/A code and P-code only.

The typical signal structure of GPS signals as presented by Langley [1997] is reproduced here as figure 6.4.

![GPS signal structure](image)

**Figure 6.4 GPS signal structure**

It can be seen from the above figure that C/A code modulates the L1 carrier and P-code modulates both L2 carrier and quadrature signal of L1 carrier. Both the codes are first modulated by the navigation message data signal prior to carrier modulation. All the carrier signals are then combined before transmission.

A comparison of the signal structure and typical characteristics of both GLONASS and GPS is given in Table 6.2.
Table 6.2 Comparison of GLONASS and GPS signals

<table>
<thead>
<tr>
<th>Type of GNSS</th>
<th>GLONASS</th>
<th>GPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signalling</td>
<td>FDMA</td>
<td>CDMA</td>
</tr>
<tr>
<td>Carrier frequency</td>
<td>L1 1602.0 - 1614.94 MHz</td>
<td>1576.42 MHz</td>
</tr>
<tr>
<td></td>
<td>L2 7/9 L1</td>
<td>60/77 L1</td>
</tr>
<tr>
<td>Type of PRN code</td>
<td>Maximal Length Sequence</td>
<td>Gold code</td>
</tr>
<tr>
<td>No. of code elements</td>
<td>C/A 511</td>
<td>1023</td>
</tr>
<tr>
<td></td>
<td>P 5110000</td>
<td>2.35·10^{14}</td>
</tr>
<tr>
<td>Code rate</td>
<td>C/A 0.511 Mbit/sec</td>
<td>1.023 Mbit/sec</td>
</tr>
<tr>
<td></td>
<td>P 6.11 Mbit/sec</td>
<td>10.23 Mbit/sec</td>
</tr>
<tr>
<td>Cross-correlation interference</td>
<td>-48 dB</td>
<td>-21.6 dB</td>
</tr>
<tr>
<td>Navigation message</td>
<td>Rate 50 bit/sec</td>
<td>50 bit/sec</td>
</tr>
<tr>
<td></td>
<td>Modulation</td>
<td>BPSK Manchester</td>
</tr>
<tr>
<td></td>
<td>Total length</td>
<td>2min30sec</td>
</tr>
<tr>
<td></td>
<td>Subframe length</td>
<td>30sec</td>
</tr>
</tbody>
</table>

It is generally found that knowledge about the capabilities of GLONASS system is rather scarce when compared with GPS, and this could be due to the little literature published so far and hence reduced number of receivers available. Details of GPS signal generation and reception is available and a huge variety of applications with GPS systems are prevailing. Also, it is easier to procure a GPS receiver for TEC related studies. Hence this study concentrates on GPS systems and their application in ionospheric studies.

6.3.3 Mathematical representation of GPS signals

The GPS signal from the kth satellite for L1 frequency can be described mathematically as

$$S_{L1} = A_p P(t) D(t) \cos(2\pi f_{L1} t + \varphi_{L1}) + A_c C(t) D(t) \sin(2\pi f_{L1} t + \varphi_{L1}) \quad (6.1)$$

where $S_{L1}$ is the signal at L1 frequency, $A_p$ is the amplitude of the P code, $P(t) = \pm 1$ is the phase of the P code, $D(t) = \pm 1$ is the data code, $f_{L1}$ is the L1 frequency, $\varphi$ is
the initial phase, $A_c$ is the amplitude of the C/A code, $C(t) = \pm 1$ is the phase of the C/A code. The GPS signal at L2 frequency is expressed as

$$S_{L2} = A_p P(t) D(t) \cos(2\pi f_{L2} t + \varphi_{L2})$$ (6.2)

As there is only one code for the L2 frequency, the equation has only one term. Here P-code is assumed. As shown, each satellite sends three rather similar signals. Any of these can be described as the product of four terms: an amplitude $A_p$ or $A_c$; the navigation data, $D(t)$; a spread spectrum code, $P(t)$ or $C(t)$; and the radio frequency (RF) carrier $\cos(2\pi ft + \varphi)$ and $\sin(2\pi ft + \varphi)$. The codes $P(t)$ and $C(t)$ are unique to each satellite.

### 6.4 Monitoring the ionosphere with GNSS

Any radio signal traversing the Earth’s atmosphere gets affected by the plasma in the ionosphere. Plasma reduces the speed of a radio signal travelling through it, more so the denser it is and more so the lower the radio frequency. This signal delay introduces error in measurements made by the Global Positioning System (GPS). To permit correction, GPS satellites transmit on two frequencies. The difference in range measured on the two frequencies by sophisticated GPS receivers provides a measurement of TEC and in turn correction as reported by Imae et al [1990] and Yun Zhang et al [2003]. Generally greater values of TECs are encountered in the day time than at night and short-term variations occur during periods of ionospheric disturbance as reported by various studies.

The measurement of TEC with Low Earth Orbiting beacon satellites and geostationary satellites has been discussed in detail in earlier chapters. The former dataset provides meridional sweeps at approximately fixed time whereas the latter provides good time resolution at fixed latitudes. Data from GPS satellites are also used in conjunction with the above two types of data for ionospheric studies as mentioned by Lanyi and Roth [1988] and Leitinger [1997]. In both the former satellite based systems, it is not possible to get a global coverage, whereas the main advantage with a GPS system for TEC studies is the global coverage available using geodetic and regional networks. Thus simultaneous data from a chain of receivers
can be used for ionospheric tomography with GNSS data, as detailed by Kunitsyn et al [2013], Zhiazhao Liu and Yang Gao [2001] and Josep Marco Pallares et al [2005]. As the frequencies used in GPS are sufficiently high, ionospheric absorption and effects of Earth’s magnetic field on the radio signals are small, and the dual frequency delays from which TECs are extracted are available in standard data formats. Also as the signals are spread spectrum modulated, GPS data provides highly accurate TEC measurements and with dual frequency measurement data, the phase ambiguity is also addressed to an extent.

However, the dual frequency GPS code and carrier phase observables are affected by several factors such as ionospheric delay, tropospheric delay, satellite and receiver clocks, instrumental biases of satellite and receiver, receiver measurement noise and multipath as explained by Ciraolo et al [2007]. Among these, the tropospheric delay, satellite and receiver clock errors are frequency independent and affect the code and carrier in the same manner and hence they can be removed easily. One method to separate the ionospheric delay from all the other components inherent in GNSS measurements is to exploit the characteristics of group versus phase delay.

When the dual frequency GPS data traverses the dispersive ionosphere, it causes a phase advance and a group delay to radio signals which is dependent on the frequency of the wave. When both the frequencies are monitored simultaneously, the data can be used in two different ways to obtain information about the TEC between any satellite and receiver as examined by Imae et al [1990]. The first way is to look at the difference between the locked signal phases which yields the change in slant TEC. The second method uses the codes on the two signals to find the difference in propagation speeds of the two signals and hence the slant TEC. Once noise and multipath have been considered, this method can ideally provide absolute TEC.
6.5 Simulation of a spread spectrum system compatible with GPS

The GPS signals are basically spread spectrum type signals with multi-carrier (dual – L1 and L2) and multi-code (dual codes – P and C/A) as explained above. As there are more satellites involved in the system, orthogonality of codes is maintained between satellites and not exactly between the ones in the same satellite. But the frequencies of the codes are made different (1.023 MHz and 10.23 MHz) so that mutual interference is reduced and in fact, in the receiver properly designed filters can take care of the decoding accurately. In order to understand the advantages of such a system, simulation software is developed in LabVIEW. The functional schematic of a multi-carrier orthogonal coded spread spectrum system is shown below in figure 6.5.

Figure 6.5 (a) Functional schematic of MC-OC SS transmitter

Figure 6.5 (b) Functional schematic of MC-OC SS receiver
The functional schematic shows two carrier signals being modulated by two orthogonal codes. One of the carrier signals is split into quadrature signals similar to L1 frequency of GPS. The other carrier is retained in single phase only. All the codes and carriers are generated using the same control signal for frequency and number of samples as explained in Chapter 4. A pair of orthogonal codes is used for simulation in lieu of the different code frequencies used in GPS systems. A Rician fading channel is assumed for ionospheric delay representation. The number of samples for the carriers and codes are considered to be the same in this simulation example. The front panel of the simulation software when run in continuous mode is shown in figure 6.6 (a) and the setting of the signals used is shown in figure 6.6 (b).

Figure 6.6 (a) Front panel of the waveforms obtained during simulation
Two different non-multiple carrier frequencies are chosen for the simulation. A pair of orthogonal MLS sequences is used for the coding purpose, whose autocorrelation is shown in the front panel. The channel shift shown as a control parameter for the simulation is indicative of the channel delay introduced during signal propagation. Provision for giving an extra delay to one of the frequencies, here f2, is also implemented in the software. This can be used to denote the differing phase delays for the different frequencies. In the receiver section, the signals are passed through fifth order band pass filters with bandwidth of $f_0 \pm 0.1$ Hz at the chosen frequencies of the carrier waves, and by the same cross-correlation method detailed in Chapter 4, it is found that the signals are recovered back. The frequencies and the code orders are varied during simulation. The channel delay is also varied. It is found that the signal recovery is ascertained to within 5% error during these changes, which indicates that usage of dual carrier and dual codes ensure unambiguous phase recovery.

Thus the hardware implementation of such a spread spectrum beacon transmitter-receiver system as used in GPS should ideally be able to recover phase unambiguously when used as a single satellite-receiver system. The details of the present GPS system used is given in the subsequent sections.
6.6 **Hardware features of a GPS system**

A GPS system consists of GPS transmitters and a ground based GPS receiver system which can receive and process the transmitted GPS signals to calculate TEC. To make it concise, the major details of each of these are described below.

**6.6.1 GPS transmitter**

In the GPS satellites, the C/A and P codes are generated precisely aligned with the clock in the satellite. In the receiver, a replica of the code is generated precisely aligned with the receiver clock, which is correlated with the received signal dominated by noise. The usage of spread spectrum codes for P and C/A codes categorize the GPS applications as one of Code Division Multiple Access (CDMA) applications. It can be considered that the satellite network uses a CDMA spread-spectrum technique where the low-bit rate message data is encoded with a high-rate pseudo-random (PRN) sequence that is different for each satellite.

**6.6.1.1 Code selection for transmitter**

It is known that a proper selection of codes help minimize CDMA noise. The code selection process is very tedious, depending on the number of codes needed and the number of codes available in a given code set. Code balance, autocorrelation side lobe peak, cross- correlation peak, and a spectral line distribution are some of the measures used in the code selection process, and these measures are not necessarily independent.

Most of the common characteristics of PN codes are stated in Chapter 2, of which the one pertinent to GPS is the ‘Code Balance’, as detailed by Spilker [1977, 1978] and Kaplan and Hegarty [2006]. A code is said to be in perfect balance if the number of 1’s is equal to the number of 0’s. For maximal length code, perfect balance is not possible because the code length is always an odd number of bits. In this case a code is said to be in balance when the difference in the number of 1’s and 0’s is 1.
There are two ways balance value can be calculated.

a. \( N_{\text{balance}} = \text{Magnitude of the difference between the number of 1's and 0's} \).

b. \( N_{\text{balance}} = \text{Magnitude of the difference between the number of bits in } \frac{1}{2} \text{ code period and the number of 1's.} \)

These two definitions differ by a factor equal to 2. Typically in the spread spectrum communication systems literature the first definition (a) is used, and in some GPS literature the second definition (b) is used. Based on the second definition a value of 6 was set as the threshold criterion for GPS systems which gives 950 codes approximately in a set of over 8000 codes satisfying the criterion.

The high autocorrelation peak and low cross-correlation peaks of C/A codes form their next important characteristic. In order to detect the presence of a weak signal, the peak of the autocorrelation of the weak signal must be stronger than the cross-correlation peaks of the strong signals. Theoretically if the codes are orthogonal, the cross correlation values will be zero. The high correlation peak property of the autocorrelation function is used to synchronize the receiver locally generated code with the code of the received signal. It can be seen that the correlation peak value is much more significant than any other correlation values contributing to secondary peaks, also called autocorrelation side lobes. Such an autocorrelation function of a typical GPS signal is shown in figure 6.7.

**Figure 6.7** Autocorrelation function with side lobes and cross-correlation function of typical GPS C/A codes
It can be seen that the peaks in the autocorrelation are at least approximately 24 dB down from the original peak so that the codes provide range measurements with high precision and without ambiguity. Also, the sharpness of the main peak helps to minimize measurement errors due to multipath. It is also seen from the above figure that the GPS codes also have low cross-correlation with each other. With uniformly weak cross-correlation, the signals from the different GPS satellites can usually be distinguished by their codes, and so the satellites broadcast simultaneously without any time sharing or offset in their transmission frequencies.

6.6.2 GPS signal generation

GPS was envisaged initially for military purposes and in order to prevent unauthorized users from using or potentially interfering with the military signal, the P-code was encrypted/modulated with the W-code, a special encryption sequence, to generate the Y-code. The Y-code is what the satellites have been transmitting and this encrypted signal is referred to as the P(Y)-code though it is generally referred to as simply P-code.

A block diagram representative of the generation of satellite signal structure for L1 and L2 is shown in figure 6.8. Here, \( f_0 \) is the fundamental frequency of 10.23 MHz so that \( L_1 = 154 f_0 \) and \( L_2 = 120 f_0 \).
Figure 6.8 Block diagram of GPS signal generation at transmitter

As shown above, the same 50-bps navigation message data is combined with both the C/A code and the P code prior to modulation with the L1 carrier. Since the C/A code data and P code data are both synchronous operations, the bit transition rate cannot exceed the chipping rate of the PRN codes, using BPSK modulation for the carrier signals. The P code data is modulated in phase quadrature with the C/A code data on L1. As shown above, the L1 carrier is phase shifted 90° before being BPSK modulated by the C/A code. Then this result is combined with the attenuated output of the BPSK modulation of L1 by the P code. The L2 frequency (1227.60 MHz) can be modulated by either the P code data or the C/A code data or by the P code alone as selected by the CS (Ground based Control segment which monitors and controls the satellite navigation). The P code and C/A codes are never present simultaneously on L2 prior to GPS modernization unlike the case with L1. In general, the P code data is the one selected by the CS.

The C/A code is a 1,023 bit pseudorandom binary sequence having good autocorrelation and cross-correlation so that the PRN code of any one satellite does not correlate well with any other satellite's PRN code. The PRN code of P-code is $6.1871 \times 10^{12}$ bits long ($6,187,100,000,000$ bits, $\sim$720.213 gigabytes). The extreme
length of the P-code increases its correlation gain and eliminates any range ambiguity associated with L1. The C/A PRNs are unique for each satellite and the P-code PRN is a small segment of a master P-code.

The navigation message gives details on each satellite's position and the network. This data modulation is done by modulo-2 addition as shown in figure 6.8. The navigation message is made up of three major components. The first part contains the GPS date and time, plus the satellite's status and an indication of its health. The second part contains orbital information called ephemeris data and allows the receiver to calculate the position of the satellite. The third part, called the almanac, contains information and status concerning all the satellites, their locations and PRN numbers.

Figure 6.9 depicts a high-level block diagram of the PRN code generation used for generation of the GPS C/A code and P code implementing the CDMA technique. Each synthesized PRN code is derived from two other code generators. In each case, the second code generator output is delayed with respect to the first before their outputs are combined by an exclusive-or circuit. The amount of delay is different for each satellite. In the case of P code, the integer delay in P-chips is identical to the PRN number. For C/A code, the delay is unique to each satellite, so there is only a table lookup relationship to the PRN number.

![Figure 6.9 PRN code generators for GPS transmitter](image-url)
The P code is bi-phase modulated at 10.23 MHz and so the main lobe of the spectrum is 20.46 MHz wide from null-to-null. The code is generated from two PRN codes with the same chip rate of 10.23 MHz. One PRN sequence has 15,345,000 chips, which has a period of 1.5 seconds, the other one has 15,345,037 chips, and the difference is 37 chips. The code length generated by these two codes is 23,017,556.5 seconds, which is slightly longer than 38 weeks. However, the actual length of the P code used is 1 week as the code is reset every week. This 38-week-long code can be divided into 37 different P codes and each satellite can use a different portion of the code. The first 32 sets of codes are used for the satellites in orbit. Five of the P code signals (33-37) are reserved for other uses such as ground transmission. In order to perform acquisition on the signal, the time of the week information must be known very accurately which can be obtained through the C/A code. A detailed block diagram of the shift register architecture is available in Kaplan and Hegarty [2006].

The C/A code is also a bi-phase modulated signal with a chip rate of 1.023 MHz and a null to-null bandwidth of the main lobe of the signal spectrum as 2.046 MHz. Each chip period is 977.5 ns (1/1.023 MHz). However, the transmission bandwidth of the GPS signal at the L1 frequency is approximately 20 MHz to accommodate the bandwidth needed for the P code signal. Therefore, in addition to the main lobe of the C/A code signal, several side lobes of the signal also get transmitted. This code is also a maximum-length PRN code, categorized as Gold codes and generated by two maximum-length linear feedback shift register (LFSR) of 10 stages, G1 and G2, driven by a 1.023 MHz clock. Both shift registers in G1 and G2 have 10 bits that the sequence length generated is $2^{10} - 1 = 1023$ bits.

The generator polynomials for both the codes are tabulated in Table 6.3 below. It can be seen that the GPS P code is generated using four 12-bit shift registers designated X1A, X1B, X2A, and X2B. The C/A code is generated by two 10-bit shift registers G1 and G2. It is also defined that all zero state is illegal for all the registers.
Table 6.3 GPS code generator polynomials and initial states

<table>
<thead>
<tr>
<th>Register</th>
<th>Polynomial</th>
<th>Initial state</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/A code G1</td>
<td>$1+X^3+X^{10}$</td>
<td>11111111111</td>
</tr>
<tr>
<td>C/A code G2</td>
<td>$1+X^2+X^3+X^6+X^8+X^9+X^{10}$</td>
<td>11111111111</td>
</tr>
<tr>
<td>P code X1A</td>
<td>$1+X^6+X^8+X^{11}+X^{12}$</td>
<td>0010010010000</td>
</tr>
<tr>
<td>P code X1B</td>
<td>$1+X^1+X^2+X^5+X^8+X^{10}+X^{11}+X^{12}$</td>
<td>01010101010000</td>
</tr>
<tr>
<td>P code X2A</td>
<td>$1+X^1+X^3+X^4+X^5+X^7+X^8+X^9+X^{10}+X^{11}+X^{12}$</td>
<td>100100100100100</td>
</tr>
<tr>
<td>P code X2B</td>
<td>$1+X^2+X^3+X^4+X^8+X^9+X^{12}$</td>
<td>01010101010000</td>
</tr>
</tbody>
</table>

A graphical representation of a typical C/A code generator is explained below as figure 6.10. The output of the G2 LFSR for each C/A code is delayed by the modulo-2 addition of two code phase selection bits specific for each satellite. The C/A code is generated by the modulo-2 addition of the output of the G1 LFSR and the delayed output of the G2 LFSR. The positions of the code selection bits determine the satellite identification.

![C/A code generator diagram](image)

**Figure 6.10** C/A code generator
Each C/A code PRN number is associated with the two tap positions on G2. An exhaustive table listed in Kaplan and Hegarty [2006] gives the tap combinations for all defined GPS PRN numbers indicative of the code phase assignments and specifies the equivalent direct sequence delay in C/A code chips. The details of these codes are needed to identify the GPS satellite that a receiver is tracking. The detail of P-code generation is also available in detail in Kaplan and Hegarty [2006].

6.6.3 GPS transmitting antenna

In order to generate uniform power over the surface of the earth, the main beam pattern of GPS transmitter onboard the satellite is of whole Earth coverage type and is achieved using an array of L band antenna elements on a panel. The transmitted signals are of right hand circularly polarized type and have gains of not less than 13.4 dB for L1 signal and 11.5 dB for L2 signals. The shaped beam pattern of the transmitting arrays are maintained to provide the required gain at the angles corresponding to the centre and near the edge of the Earth resulting in slightly increasing transmitting antenna array gain in between the nadir angles of ± 5°. For the angular range of ± 14.3 degrees from boresight, L1 ellipticity is better than 1.2 dB and L2 ellipticity better than 2.2 dB, as given in the interface control document ICD-GPS-200C [1993].

6.7 Link budget

As detailed by James Bao-Yen Tsui [2000], the minimum power levels of the signals at the receiver is better than -160 dBW for L1 C/A code, -163 dBW for L1 P code and -166 dBW for L2 P code. The link budget calculation for a digital system like spread spectrum (GPS) system involves a two-step process according to Hausman [2009] and Samir H. Abdul-Jauwad. The first step is to derive the basic SNR of the receiver, as shown for L1 C/A code in Table 6.4. Here, the effective system noise temperature includes the ambient temperature and equipment noise factors as the dominant factors, in addition to the contribution by sky noise, antenna noise temperature and line losses, as mentioned by Braasch and Dierendonck [1999]. It is
also seen that at the receiver output, only white noise is distinguishable as SNR is negative

Table 6.4 Link budget calculation for GPS L1 frequency

<table>
<thead>
<tr>
<th>Frequency (MHz)</th>
<th>1576.42 (C/A code)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tx Power O/P (W)</td>
<td>478.63W</td>
</tr>
<tr>
<td>EIRP (dBW)</td>
<td>26.8</td>
</tr>
<tr>
<td>Free space loss for max elevation (dB)</td>
<td>-182.4</td>
</tr>
<tr>
<td>Atmospheric attenuation (dB)</td>
<td>2.0</td>
</tr>
<tr>
<td>Signal Power at antenna I/p (P_s) dBW</td>
<td>-157.6</td>
</tr>
<tr>
<td>System Noise Temp (dB/°K)</td>
<td>513 K (27.10)</td>
</tr>
<tr>
<td>Boltzman Constant (dBW/Hz/°K)</td>
<td>-228.6</td>
</tr>
<tr>
<td>Noise power (dBW/Hz)</td>
<td>-201.5</td>
</tr>
<tr>
<td>IF filter bandwidth (typ.)</td>
<td>2 MHz</td>
</tr>
<tr>
<td>Effective noise power (dBW/Hz)</td>
<td>-138.5</td>
</tr>
<tr>
<td>SNR (dB)</td>
<td>-19.1</td>
</tr>
</tbody>
</table>

The second step of link calculation deals with the signal extraction after correlation and demodulation with the help of filters. The link budget is generally used to ensure that the demodulator has adequate C/N_0 (ie, carrier power to noise power spectral density ratio) for all required operating conditions.

Now, assuming a post-correlation bandwidth of 50 Hz for the GPS signals, the effective noise power in 50-Hz bandwidth becomes approximately -184.5 dBW. Thus after correlation and de-spreading, the SNR has increased to:

\[
\text{SNR} = \text{Signal power in dB} - \text{Noise power in dB} \\
= -157.6 - (-184.5) = 26.9 \text{ dB}
\]  

(6.3)
The increase in SNR as a result of de-spreading in this case is

\[ \text{SNR gain (dB)} = 26.9 - (-19.1) = 46 \text{ dB}. \]

Thus it can be seen that pre correlation SNR’s are negative whereas post correlation SNR’s are positive. It is convenient, therefore, to normalize the SNR to a 1-Hz bandwidth and thus achieve a ratio of signal and noise which is bandwidth-independent. Alternately, this can be viewed as a density and the result referred to as the carrier-to-noise density \((C/N_0)\) ratio. This is related to the SNR as

\[
\frac{C}{N_0} (dB) = 10 \log_{10}(SNR) + 10 \log_{10}(B) = \text{SNR (dB) + B (dB)}
\]

(6.4)

which is the straight ratio form of the SNR at a certain point in the receiver, like the final IF stage, and B is the bandwidth (in Hz) of that stage of the receiver. Thus in the above case, the pre and post-correlated carrier-to-noise density is calculated as

\[
\frac{C}{N_0} (\text{pre – correlation}) = -19.1 + 10 \log_{10}(2.046 \times 10^6) = 44.1 \text{ dBHz} \quad (6.5a)
\]

And

\[
\frac{C}{N_0} (\text{post – correlation}) = 26.9 + 10 \log_{10}(50) \sim 44 \text{ dBHz} \quad (6.5b)
\]

Thus it is seen that \(C/N_0\) remains constant throughout the receiver system. The actual received satellite signal power varies with user antenna gain, satellite elevation angle, and satellite age as mentioned by Braasch and Dierendonck [1999]. The typical range of \(C/N_0\) is from 35–55 dB-Hz.

The other important parameters are the ratio of energy per bit to noise density \((E_b/N_0)\) and the probability that a bit sent is received incorrectly termed as Bit Error Rate (BER), as shown by Martin Dottling and Simon Saunders [1999]. According to Jim Pearce [2000], \(E_b/N_0\) is classically defined as the ratio of Energy per Bit \((E_b)\) to the Spectral Noise Density \((N_o)\). It is measured at the input to the receiver and is used as the basic measure of how strong the signal is. The different forms of digital modulation techniques like BPSK, QPSK, QAM, etc. have different curves of theoretical bit error rates versus \(E_b/N_o\) which is reproduced here as figure 6.11. The
bit error arises due to symbol errors occurring because of thermal noise, external interference and inter-symbol interference.

![Figure 6.11 Probability of Bit error for common modulation methods](Source: Intersil)

It is known that GPS signals are transmitted with BPSK modulation which implies a change of 1 bit per symbol so that both symbol rate and bit rate are equal. Thus

\[ \frac{E_b}{N_0} = \frac{C}{N_0} + 10 \log_{10} \left( \frac{\text{Symbol rate}}{\text{Bit rate}} \right) \]

becomes

\[ \frac{E_b}{N_0} = \frac{C}{N_0} \].

As mentioned above, PRN code of C/A code is 1023 bits long spreading the signal at a rate of 1.023 Mbps. A measure of the efficiency of signal spreading is given by the parameter processing gain which is defined as,

\[ \text{Processing gain} = 10 \log \left( \frac{\text{Chip rate}}{\text{Data rate}} \right) \]  \hspace{1cm} (6.6)

and is 43.1 dB for the C/A code signal with chip rate of 1.023 Mbps and data rate of 50 bps. Thus, it can be seen that a typical spectrum of the GPS signals from the visible satellites is not apparent as it lies well below the noise floor as shown in figures 6.12 (a) and 6.12 (b). After being correlated with the locally generated code,
however, the signal is de-spread thus occupying the bandwidth of the navigation data, namely, 50 Hz as indicated in figure 6.12 (c).

**Figure 6.12 (a)** Time domain plot of raw GPS signal plus noise, sampled at approximately 5 MHz, after front-end processing (amplification and filtering only)

**Figure 6.12 (b)** Frequency spectrum of the pre-correlation FFT of the GPS signals plus noise plotted in figure 6.12 (a)

**Figure 6.12 (c)** Frequency spectrum of post-correlation FFT of the GPS C/A signal present in the data plotted in Figure 6.12 (a)
As GPS receivers suitable for ionospheric studies are readily available, it is better to arrive at the specifications pertaining to the receiver front end, so that a suitable one can be chosen. In order to check if a receiver has enough SNR to track GPS signals, let us first assume a BER of $10^{-5}$ in the baseband processor, then corresponding $\frac{E_b}{N_0}$ for the BPSK modulation pre-processor is $\sim 10.1$ dB as obtained from the graph of $\frac{E_b}{N_0}$ Vs. BER shown above in figure 6.11. The mathematical relation for generation of the graph in figure 6.12 for the BPSK modulation of GPS signals is given as

$$BER(t) = \frac{1}{2} \text{erfc} \left( \frac{E_b}{N_0} (t) \right)$$

(6. 7)

Now, SNR at the correlator input = Required ($\frac{E_b}{N_0}$) – Processing gain = -33 dB.

This differs from the SNR derived above by $\sim 14$ dB, which indicates that a receiving antenna with again of minimum 14 dB would suffice our purpose, as a 0 dB antenna has been assumed in Table 6.4 above.

**6.8 Principle of working of a GPS receiver system**

The bi-phase coded GPS signal has the carrier and code frequencies varying due to Doppler Effect, which is caused by the motion of the GPS satellite as well as from the motion of the GPS receiver. The GPS signal received from the antenna is a combination of carrier wave, C/A code and navigation message, along with this Doppler. In order to extract the navigation data from the GPS signal, it is necessary to remove the carrier wave and the C/A code. The process of receiving GPS signals may be divided into three steps: acquisition, tracking, computing the position solution from recovered navigation data bits. Acquisition is used to detect the presence of a signal from a particular satellite and calculate the initial code offset and Doppler shifted carrier frequency. The tracking process, bit synchronization, and sub-frame synchronization are used to keep lock on the carrier and code and to obtain pseudorange, carrier phase measurements from the recovered navigation data bits. In a conventional receiver, the front end, the acquisition and tracking processes, are implemented in hardware while the navigation solution calculations are
completed in software. Several designs have been reported, specifically by Louis Litwin [2001], Darius Plausinaitis [2009] and lecture notes of Tan Wong.

The two most important parameters to be determined by the acquisition program are the beginning of the C/A code and the carrier frequency of the input signal. The typical received signal contains information from more than one GPS satellite. For each satellite, the beginning of the C/A code and the carrier frequency after Doppler shift are different from others. Once the acquisition program detects the presence of a desired satellite and finds the beginning of the C/A code, it uses the information to de-spread the signal. After de-spreading the spectrum, the signal becomes a continuous carrier wave. The carrier frequency of the signal can then be determined. The estimated beginning of the C/A code and the carrier frequency are then passed onto the tracking program.

The beginning of the C/A code is found by correlating the incoming signal with the receiver generated signal. There are several different methods to perform acquisition. The two popular methods are the serial search in time domain and parallel search using fast Fourier transform (FFT) in frequency domain. Though serial search is the slowest search method, it is usually implemented in hardware based receivers due to its simplicity, as detailed by Jovanovic [1988]. The parallel search (FFT method) in frequency domain is usually implemented by software receiver since serial search method is computation intensive in the software approach.

After the acquisition process is complete, the receiver enters the tracking phase. A delay-locked loop (DLL) is used to track the C/A code phase and a phase-locked loop (PLL) is used to track the carrier frequency of the incoming signal with Doppler shift. When both tracking loops are in lock, it is possible to decode the 50 Hz navigation data message. The tracking loop follows the incoming signal and adjusts itself to de-spread and de-modulate the incoming signal. If the receiver is stationary, the rate of change of Doppler frequency is small so the update rate of tracking loop is less frequent. Since the phase of the incoming GPS signal is not
constant, the code tracking process becomes necessary to remove the phase shift in the C/A code such as the effect of noise.

These two phase locked loops can be coupled together as shown in the block diagram of figure 6.13. Here, the C/A code loop generates three outputs: an early code, a late code, and a prompt code. The prompt code is applied to the digitized input signal and strips the C/A code from the input signal. Stripping the C/A code means to multiply the C/A code to the input signal with the proper phase. The output is a continuous signal with phase transition caused only by the navigation data. This signal is applied to the input of the carrier loop. The output from the carrier loop is also a continuous signal with the carrier frequency of the input signal. This signal is used to strip the carrier from the digitized input signal giving an output signal with only a C/A code and no carrier frequency, which is applied to the input of the code loop.

**Figure 6.13** Code and carrier tracking loops
The acquisition program determines the beginning of the C/A code. The code loop generates early and late C/A codes and these two codes are the C/A code time shifted typically by approximately one-half-chip time of 0.489 μsec or less i.e., positioned in time at ±0.5 chip. The early and late codes correlate with the input C/A codes to produce two outputs. Each output passes through a moving average (MA) filter and the output of the filter is squared. The two squared outputs are compared to generate a control signal to adjust the rate of the locally generated C/A code to match the C/A code of the input signal. The one with the highest correlation value is selected and retained. This locally generated C/A code is the prompt C/A code and this signal is used to strip the C/A code from the digitized input signal.

The carrier frequency loop receives a CW signal phase modulated only by the navigation data as the C/A code is stripped off from the input signal. The acquisition program determines the initial value of the carrier frequency. The voltage-controlled oscillator (VCO) generates a carrier frequency according to the value obtained from the acquisition program. This signal is divided into two paths: a direct one and one with a 90-degree phase shift. These two signals are correlated with the input signal. The outputs of the correlator are filtered and their phases are compared against each other through an arctangent comparator. The output of the comparator is filtered again and generates a control signal. This control signal is used to tune the oscillator to generate a carrier frequency to follow the input CW signal, so that the loop can continuously demodulate the incoming signal according to Parkinson and Spilker [1996] and James Bao [2000]. The PLL used to track a GPS signal is usually a second order loop. However, a higher order PLL is necessary for tracking signals on a high dynamics receiver.

The arctangent discriminator is insensitive to the phase transition caused by the navigation data and it can be considered as one type of a Costas loop. A Costas loop is a type of PLL which is insensitive to phase transition of the navigation data bit transitions. This is used to determine the phase shift in the carrier frequency. Once the input IF signals in the $I$ and $Q$ channels are down-converted to baseband using the estimated carrier frequency, the resulting baseband signals are then multiplied with the prompt code and are summed over one code period. The resulting $I$ and $Q$
signals are then fed into the discriminator which gives the angle in radians between the signals in the two channels.

Data demodulation is then done by despreading the 1.023 M chip/sec to 1000 bits/sec bit stream and with a bit synchronizer to recover the 50 bits/sec information. For bit synchronization, the beginning of a bit needs to be identified first. This is done by finding the zero crossing edge (at 01), which indicates the beginning of a bit. Then 1000 bits/sec is partitioned to 20 msec intervals and bit samples in a 20 msec are summed and averaged to decode the navigation data.

6.9 Generic GPS receiver system

A high-level block diagram of a modern generic digital GPS receiver is shown in figure 6.14 and its various blocks are explained below.

![Block diagram of a generic GPS receiver](image)

**Figure 6.14** Block diagram of a generic GPS receiver

The GPS RF signals of all satellites in view are received by a RHCP antenna with nearly hemispherical (i.e., above the local horizon) gain coverage. This pattern allows tracking of satellites from zenith almost down to the horizon for all azimuths.
Typical coverage is $160^\circ$ with gain variations from about 2.5 dBiC at zenith to near unity at an elevation angle of $15^\circ$. Below $15^\circ$, the gain is usually negative. A typical antenna pattern is shown in figure 6.15. This pattern was produced by a stacked patch antenna element in a dielectric substrate. This particular antenna is designed to operate at both L1 and L2, but only the L1 pattern is given here for illustration.

![Antenna Pattern](image)

**Figure 6.15** Example of RHCP antenna pattern (*Source: Kaplan & Hegarty*)

The antenna is designed for the frequency of reception of the receiver system. Some of the characteristic technical parameters that contribute to the selection of antennas for GPS applications include the number of frequencies to be tracked, reception bandwidth, gain characteristics, multipath rejection capability, active or passive, phase centre stability, efficiency, form factor and transfer response. Of particular importance is multipath rejection capability and phase centre stability as explained by Kim *et al* [2004], Gilad *et al* [2008] and Wu *et al* [2008]. Multipath is the phenomenon whereby a signal arrives at a receiver site via two or more different paths. The difference in path lengths causes the signals to interfere at the receiver. This can be minimized by effective antenna beam shaping as reported by David Wells [1987]. In the transfer response, it is desirable for the magnitude response to be nearly constant as a function of frequency and for the phase response to be linear with frequency within the pass band of interest. The antenna phase response
determines the phase centre of the antenna. The phase centre of an antenna structure is defined as the apparent source of radiation. This in turn is responsible for the accuracy of phase recovered by the receiver. The phase centre of the antenna is thus not constant but is dependent upon the observation angle, i.e., it is angle dependent. In modern versions of GPS receivers, the optimization of antenna phase center dependency is made part of the receiver processing software.

The RF signals from the antenna in figure 6.14 are amplified by a low noise preamplifier (preamp), which effectively sets the noise figure of the receiver and typically has a gain on the order of 25–40 dB. This LNA may be embedded in the antenna housing (or radome) itself in some cases which helps to maintain a low-noise figure within the receiver. Such an antenna requires power, which is usually supplied by the receiver front end through the RF coaxial cable. LNAs have a typical noise figure less than 2 dB, but the addition of pre-selection filtering, burnout protection, and other associated losses usually result in an overall noise figure of 3–4 dB, according to Braasch and Dierendonck [1999]. The antenna and receiver front end must have sufficient bandwidth to pass the signals of interest. This is considered to be within 1% to 2% of the center frequency. It should be noted that the receiver’s antenna/front-end bandwidth is directly proportional to the accuracy required for the specific application of the receiver. That is, the more frequency content of the received satellite signal that is processed, the better will be the accuracy performance.

The RF susceptibility characteristics of GNSS receivers to adjacent-band RF interference must be considered when evaluating the overall effectiveness of the system for any potential application. Only front-end filtering or antenna cancellation techniques can prevent the GNSS receiver from being overdriven by adjacent out-of-band signals. Thus there can be a passive bandpass filter between the antenna and preamp to minimize out-of-band RF interference.
The amplified and signal conditioned RF signals are then down-converted to IF using mixing frequencies from local oscillators (LOs). The LOs are derived from the reference oscillator by the frequency synthesizer, based on the frequency plan of the receiver design. One LO per down converter stage is required. Two-stage down-conversion to IF is typical, but one-stage down-conversion and even direct L-band digital sampling are also used. However, since nearly 100 dB of signal gain is required prior to digitization, placing all of this gain at L-band is conducive to self-jamming in the receiver front end, so down-conversion is the best choice. The LO signal mixing process generates both upper and lower sidebands of the satellite signals. The lower sidebands are selected and the upper sidebands and leak-through signals are rejected by a post mixer bandpass filter. One advantage of multistage down conversion is the opportunity to isolate the gain stages and to use increasingly higher-Q filters as the IF is lowered. The preferred conventional filter for GNSS at lower IF stages have maximally flat in-band and sharp stop-band roll-off characteristics similar to Chebyshev band pass filter.

The signal Doppler and the PRN codes are preserved after the mixing process. Only the carrier frequency is lowered, but the Doppler remains referenced to the original L-band signal. These are then converted to digital domain using ADCs, which are phase locked to the reference oscillator in the receiver system. The IF must be high enough to provide a single-sided bandwidth that will support the PRN code preferred in most cases since the signals from all GPS satellites in view are buried in thermal noise at IF, as is understood from the SNR calculation explained above.

The digitized IF signals are ready to be processed by each of the $N$ digital receiver channels. No demodulation has taken place, only signal gain and conditioning plus A/D conversion into the digital IF. The digital receiver channel block is designed with DSP/FPGA and usually has the baseband functions such as the loop discriminators and filters, data demodulators, SNR and phase lock indicators.
6.10 GPS system chosen for measurement of TEC

From the study of the different receiver systems available for GPS TEC measurements, the one most suited for the present study is identified as Model Propak-V3 shown in figure 6.16 (a) with antenna GPS 702 GG of M/s. Novatel, pictorially shown in figures 6.16 (b) and 6.16 (c).

![Figure 6.16 (a) Dual channel GPS receiver chosen for TEC measurement](image)

This antenna is a high-performance L1/L2 choke ring antenna which substantially reduces the effects of multipath, making it ideal for TEC measurement applications. A choke ring antenna is a particular form of omnidirectional antenna for use at high frequencies. It is fabricated by constructing concentric ring structure called corrugation around a small aperture antenna like crossed dipole, horn or patch antenna. The choke ring depth which is a function of the wavelength of reception plays a significant role in achieving controllable and repeatable antenna characteristics like phase centre stability and radiation pattern. Due to its delicate
construction, it is often enclosed in a protective cover or radome. Choke ring antennas are notable for their ability to reject multipath signals from a source. Since the path that a signal takes from a transmitter to receiver can be used to measure the distance between the two, it is highly suited for GPS and radar applications. A single choke ring antennas can be designed for reception of both L1 and L2 frequencies. Low-elevation angle signals are nulled by the rings, and thus the antenna’s gain effectively is reduced at these low angles as detailed by Kalyanaraman et al [2006].

The present antennas are designed to receive signals from both GPS and GLONASS satellite systems, and they include patented Pinwheel technology of M/s. Novatel to provide multipath rejection in a compact and lightweight antenna. A highly stable antenna phase center makes the Model 702-GG (L1 and L2) antennas the perfect choice for high precision applications. The antenna is waterproof to IEC 60529 IPX7 and meets the MIL-STD-810F specification for vibration and salt spray, resulting in an antenna suitable for adverse conditions.

The receiver system ProPak-V3 shown in figure 6.16(a) is a durable, high-performance receiver with 72 available channels including GPS L1, L2 and GLONASS. It has a measurement precision of 6 cm on L1 C/A code and 0.75 mm on L1 Carrier Phase while for L2 P code, this is 25 cm with a carrier phase resolution of 2 mm, when operated in differential mode. The receiver has a time accuracy of 720 nsec and a signal reacquisition time of better than 1 sec in both the carrier frequencies. It takes +9 to +18 VDC power and provides + 5V DC as the antenna power, taken through the antenna TNC connector. The receiver communicates to PC through a RS-232 serial port capable of a maximum baud rate of 230,400 bps, though it is configured for 9600 bps by default. Provision for external control oscillator is also provided in the back panel. This can be used to take care of the oscillator drift in the receiver unit to an extent by connecting a highly stable oscillator like Rubidium based. The receiver supports various data types including NMEA 0183 version 3.01 which is the one used in our present application, some of the details of which are mentioned in Chapter 4. The receiver also provides a configurable 1 PPS output, which can be used for time synchronization of other instruments in lab.
When the receiver is switched on first, in order to determine which satellites are visible and which constellation of visible satellites is the most suitable, three things are needed:

- An up-to-date almanac
- Rough estimates of user position and velocity
- An estimate of user GPS time.

If any of these parameters are missing or obsolete, the receiver performs a sky search. If all are available, then using the user position, the GPS time estimate, and the almanac, the satellite positions and LOS Doppler can be computed. Using the estimated user position and the satellite positions, the visible Satellite Vehicles (SVs) can be determined. From the list of visible SVs and the user position, typically the best constellation geometry for good dilution of precision is determined. When the constellation has been selected, the search process begins. From the user velocity and the satellite LOS Doppler, the total LOS Doppler can be determined. This is used in the Doppler search pattern for the satellite. If the approximate time and position are known and the ephemeris data has been obtained during a recent previous operation, the time to first fix can be around 30 seconds for a typical multichannel GPS receiver if the signals are unobstructed. If the ephemeris is not available for the first fix, the almanac data is used until more precise data become available. Once the GPS receiver has performed this sky search and gets locked to the satellite constellation, the receiver starts acquiring data which can be used for the specific application.

### 6.10.1 Mathematical representation of TEC calculations performed in the GPS receiver

The refractive index of ionosphere when an electromagnetic wave passes through cold magnetized plasma as given by Appleton-Hartree equation from Davies [1989] is

\[
\xi(x) = 1 - \frac{1}{2} X \pm \frac{1}{2} i Y \cos \theta \left[ 1 \mp \frac{1}{4} (1 + \cos^2 \theta) \right]
\]

(6.8)

where

\[
X = \frac{N_e e^2}{4\pi^2 \varepsilon_0 m f^2}, \quad Y = \frac{B_0 |e|}{2\pi m f}
\]
and \( \theta \) is the angle between the direction of signal propagation and the magnetic field.

For L-band GPS signals, the following expression is a good approximation for index of ionosphere, as given by Parkinson and Spilker [1996].

Assuming \( N=10^{12} \), \( B_0 =0.5 \times 10^{-4} \), \( \theta=0 \), the idealized refractive index is

\[
n = 1 - \frac{\alpha N}{f^2}
\]

(6.9)

where \( \alpha = \frac{e^2}{8\pi^2 m\epsilon_0} \approx 40.3 \)

This does not consider satellite clock errors, special and general relativistic terms, phase ambiguity bias, random noise. From this, the phase advance and group delay can be derived as

\[
\rho_\theta = \int_s (n dS) = \int_s \left[ 1 - \frac{\alpha N}{f^2} \right] dS = \rho - I
\]

(6.10)

and \( \rho_p = \int_s (n dS) = \int_s \left[ 1 + \frac{\alpha N}{f^2} \right] dS = \rho + I \)

(6.11)

where \( I = \frac{\alpha}{f^2} \int N dS \approx 40.3 \frac{TEC}{f^2} \), and is also known as range error.

Thus, for carrier phase measurement the minus sign is used while for pseudorange measurement the positive correction is applied.

When there are two frequencies, as in the case of GPS signals, it is assumed that both the frequencies follow the same ionospheric path, so that the TEC integration path is same and higher order ionospheric terms as detailed in Davies [1989] and Shkaforsky [1961] are ignored. This has a basic disadvantage that the systematic bias becomes insufficient when ionospheric bending is large, i.e., when the slant range between the satellite and ground receiver is high. After getting the pseudo ranges at each frequency from the above equations, the TEC for two coherent frequencies is calculated as

\[
TEC = \frac{f_1^2 f_2^2 (R_1 - R_2)}{40.3 (f_2^2 - f_1^2)}
\]

(6.12)

where \( R_i \) and \( f_i \) \((i = 1,2)\) are the pseudorange and the corresponding frequency, respectively.
The TEC obtained from the above equations form the delays of radio signals on channels L1 (1576.42 MHz) and L2 (1227.6 MHz) and is termed as oblique or slant total electron content (STEC) of the ionosphere as defined by Hofmann-Wellenhof et al [2001]. The measured delay contains the delays in the ionosphere, the satellite transmitting system and the receiving system including antennas, cables etc.

### 6.10.2 Data acquisition and processing

The system is connected to the RS 232 serial port of the PC and powered on. Software in C++ is developed for parsing and continuous logging of the data from the GPS receiver. The program generates an output header file which gives details of the week number, time in milliseconds, PRN, azimuth and elevation of the satellites being tracked. The science data file gives the various parameters like TEC, VTEC and Scintillation index S4 derived from received data. The data file also has details on UT, sub-ionospheric latitude, sub-ionospheric longitude, satellite azimuth and elevation, satellite PRN number and amplitude value of L1 signal every second. The software is programmed to collect data for 24 hour duration, save onto a single data file and start a new one immediately.

The inbuilt software in the receiver estimates the STEC over each receiving GPS station by means of a Kalman filter as explained by Sardon et al [1994] during a 24-hour run for each satellite-receiver combination, assuming a second order polynomial approximation. Also, it is assumed that during this 24 hour, the biases are assumed to be constant. This STEC estimation procedure reveals TEC data measured along permanently changing satellite links. To obtain normalized data, the slant TEC data is converted to the vertical TEC data by a mapping function $M(\varepsilon)$ which is in general defined by:

$$M(\varepsilon) = \frac{\text{STEC}}{\text{VTEC}}$$ (6.13)

For a single-layer approximation of the ionosphere, we obtain:

$$M(\varepsilon) = \frac{1}{\sqrt{1 - \left(\frac{r_E \cos \varepsilon}{r_E + h_{sp}}\right)^2}}$$ (6.14)
where \( \varepsilon \) is the elevation angle, \( r_E \) is the Earth radius and \( h_{sp} \) is the height of the assumed ionospheric layer at the sub-ionospheric point SP. Taking into account simulation calculations with realistic electron density profiles, the height \( h_{sp} \) of the ionospheric shell is taken as \( h_{sp} = 350 \text{ km} \) as detailed by Ramarao et al [2006]. Thus vertical electron content (VTEC) data at the sub-ionospheric points is calculated which can be used for further scientific analysis and generation of regional maps, following Jakowski et al [1996]. Because of the uncertainties on the receiver and transmitter delays, antenna characteristics, ionospheric gradients and thin shell approximation of ionospheric height for conversion of slant TEC to vertical TEC, the overall accuracy of the present GPS TECs is not better than 3-4 TEC units.

It is possible to discard the lower elevation angles by entering the appropriate angle during data acquisition itself. The processed data get stored in the filename starting with year and Julian day. Eg. The filename 2007008 indicates that the file is from year 2007 and day number is 8, which indicates it as January 8.

**6.10.3 Data interpretation**

A sample data file for a typical day, 01.12.2007, processed with the GPS system installed at SPL, VSSC is shown in the figures below. Figure 6.17 shows the STEC and VTEC plots for the entire day. The X-axis shows the System Local time (SLT) from 0:00 hrs to 24:00 hrs. The blue curve shows the STEC measured and the green curve shows the calculated VTEC, assuming the sub-ionospheric height of 400 km according to the formula given in equation 6.14 above. It can be seen that both the graphs follow almost the same trend. Also, it is seen that the both curves are not single line curves as is observed in the case of LEO and GSAT beacon systems referred in earlier chapters, and that they have a range of variation at any instant. This is because the receiver tracks more than one satellite at any given instant depending on the satellite elevation. The output data file has TEC samples every 30 seconds approximately, so that in 2 minutes, we have data values from at least 3 satellites which are of high elevation. The black curves in the figure shows the trend line, calculated with a moving average of 20 samples.
The calculation of VTEC is dependent on the satellite azimuth and elevation above the location, and hence the VTEC will not be exactly following the STEC curve. This can be explained with the following figures. Figure 6.18 (a) shows a case where the satellite pass is almost overhead, as can be seen from the zenith plot. Here the satellite is moving from 6.11°N, 76.44°E to 11.67°N, 77.95°E. The VTEC calculated depends on the observation angle \( \chi \), explained in Chapter 3. So when the satellite is almost overhead, this angle \( \chi \) is zero, indicating almost equal values for STEC and VTEC. Once the satellite moves away, the angle \( \chi \) increases giving a corresponding deviation between the TECs. In figure 6.18 (b), the satellite footprint is from 10.17°N, 79.53°E to 10.07°N, 79.57°E meaning the satellite is moving near the horizon without any overhead duration. This indicates that the \( \chi \) value is almost constant as can be seen from the zenith plot also. In this case, it can be seen that VTEC almost follows STEC.
Figure 6.18 (a)

Figure 6.18 (b)

Figure 6.18 Difference in VTEC and STEC for differing satellite elevation angles
The variation of the satellite zenith angle for all the satellites for 01.12.2007 over Trivandrum is shown in figure 6.19. The receiver tracks all the satellites from 5°N to 12°N latitude and longitude between 73.97°E to 80.22°E. The geographic location of Trivandrum is 8.5°N, 76.9°E; so that an antenna with an almost hemispherical pattern as indicated in figure 6.15 will be able to have the above tracking range. Here, in the particular day used for case study, satellites with elevation angles above 35 degrees only are tracked, as the acquisition software itself removes the lower elevation passes.

**Figure 6.19** Variation of satellite zenith angle for the passes over Trivandrum

In order to understand the TEC variation further, the data for the satellite passes having almost equal elevation and same local time is identified and divided into blocks of 4 hour each, starting from 0:00 hours. The graph in figure 6.20 (a) shows
the TECs obtained from three satellites for the first 4 hour block i.e., 0:00 to 4:00 hours. The satellite PRN number is listed on the right side. It is clearly visible that none of the TEC curves overlap, indicating that the TEC derived from all the cases differs, though all are being obtained over the same ground location. Similar deviation observed for two more datasets is shown as figures 6.20 (b) and 6.20 (c). The satellite PRN 2 and 12 in figure 6.20 (b) have maximum elevation of 70.5° and 72.5° and are tracked by the receiver from 1200 to 1600 hrs. The satellite PRN 14 and 21 in figure 6.20 (c) have maximum elevation of 62.65° and 63.2° and are observed between 2000 hrs and 2400 hrs. It can be seen that all these plots convey similar observational results irrespective of the time of the day. Similar observations have been reported earlier from other locations also. However, these variations are not constant and show diurnal and seasonal variation as well as solar dependency.

Figure 6.20 (a)

![Figure 6.20 (a)](image1)

Figure 6.20 (b)

![Figure 6.20 (b)](image2)
This difference observed indicates that there is an uncertainty in the measurement of VTEC using GPS satellites, which is mainly contributed by the bias errors.

Thus even though TEC measurement with GPS is highly precise, it is often rendered inaccurate mainly due to satellite and receiver differential code biases (DCBs) or instrumental biases. Calculated satellite DCB values are now available from a variety of sources, but receiver DCBs generally remain an undertaking of receiver operators and processing centres. The topic of IFB determination has been addressed before, notably by Mannucci et al [1998], Coco et al [1991], and Sardon and Zarraoa [1997], Bilitza et al [2003], Richard Dear and Cathryn Mitchell [2006] and Cornely [2013]. It has been observed that the magnitudes of these biases and the accuracy to which they can be determined is very important for ionospheric measurement and further tomographic imaging, since they are a source of error for ionospheric measurements based upon differential code records. The majority of ionospheric research using GPS data uses a mapping technique to approximate the ionosphere as a thin shell of ionization at a nominal height of the ionosphere, usually between 350 and 450 km. However, Birch et al [2002] have demonstrated that an inappropriate choice of shell height is also another major source of error in TEC determination.

However, to account for the effects of instrumental biases, the TEC equation can be modified as
Chapter-6

\[
\text{TEC} = \frac{f_1^2 f_2^2 (R_1 - R_2 - b_{rx} - b_{sv})}{40.3 (f_2^2 - f_1^2)}
\]

\[= -9.52 \times 10^{16} \ (R_1 - R_2 - b_{rx} - b_{sv}) \quad (6.15)\]

where \(b_{rx}\) and \(b_{sv}\) are the receiver inter-frequency bias and satellite inter-frequency bias, respectively. A first investigation on the carrier phase and code pseudorange measurement combination was provided by Hatch [1982]. Lachapelle et al [1986] gave a varied weighted algorithm of smoothing code pseudorange with carrier phase measurements, which could detect cycle slips though corrections were not implemented.

6.11 Summary

The GPS system is a spread spectrum modulated system with dual carrier (L1 and L2) and dual codes (P and C/A). A beacon system similar in principle with dual carrier and orthogonal codes is simulated for phase measurement and is found to give accurate results. The hardware details and working of a typical GPS transmitter-receiver system readily available for TEC measurement is studied. Results of a sample data file are also discussed. It is seen that when multiple satellites appear simultaneously within the receiver antenna look-angle, the inter-frequency biases comes into picture. The main challenge at this point is to separate the code inter-frequency biases (IFBs) from the line-of-sight TEC. Since both terms are linearly dependent, a mathematical representation of the TEC is usually required to obtain an estimate of each quantity. Misspecifications in the model and mapping functions are found to contribute significantly to errors in the IFB estimation, which has been an ongoing research topic for the past two decades and still remains an issue for accurate TEC determination.